The Geometry of Polynomials in Algorithms

Lecture 1: Real Rooted Polynomials

Lecturer: Nathan Klein

1 Introduction

Given a discrete probability distribution $\mu : \mathbb{Z}_{\geq 0} \to [0,1]$ with finite support, we define p_μ as its generating polynomial:

$$p_\mu(t) = \sum_{\alpha \in \mathbb{Z}_{\geq 0}} \mu(\alpha) t^\alpha$$

For example, suppose μ is 0 with probability $\frac{1}{4}$, 1 with probability $\frac{1}{2}$, and 2 otherwise. Then:

$$p_\mu(t) = \frac{1}{4} + \frac{1}{2}t + \frac{1}{4}t^2 = \left(\frac{1}{2}t + \frac{1}{2}\right)^2$$

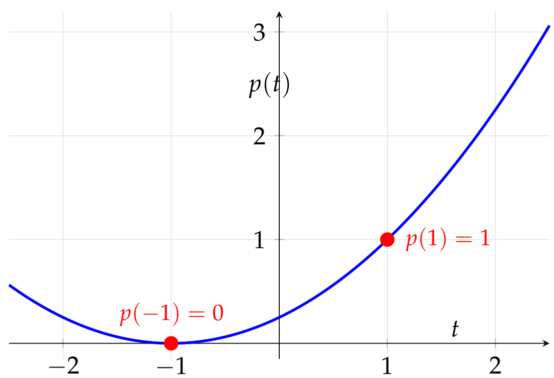

Figure 1: The encoding of the distribution of the number of heads in two independent coin tosses as a polynomial. The polynomial is degree 2 and has a double zero at -1.

At first, it may appear kind of pointless to encode μ in this way, although it's a nice way to see that our distribution factors into two independent coin tosses. You also might notice that $p_\mu(-1) = 0 = \mathbb{P}[\text{even}] - \mathbb{P}[\text{odd}]$, which is fun but not exactly revolutionary.

But let's keep exploring. What might we be able to learn about distributions that arise from independent coin tosses?

1.1 Poisson Binomial

Now let's suppose we flip d different coins with success probabilities $p_1, \dots, p_d > 0$. Let μ be the distribution counting the number of heads. This is known as a Poisson binomial distribution. The generating polynomial is:

$$p_\mu(t) = \prod_{i=1}^{d} (p_i t + (1 - p_i))$$

Reading things off: we increase the number of heads by 1, thus incrementing the power of t, with probability p_i. Otherwise we do not.

Notice that $p_\mu(1) = 1$. That's a good sanity check, it means that the probabilities sum to 1. Now, let's try to answer the following question: is a Poisson binomial distribution unimodal? In other words, does $\mathbb{P}[X=0], \mathbb{P}[X=1], \ldots, \mathbb{P}[X=d]$ (the *rank sequence* of μ) increase until a certain point and then decrease?

You can answer this question without looking at p_μ, but let's see how we can use properties of p_μ to do this in a particularly nice way which will tell us even more about the sequence of probabilities.

Three Views of a Polynomial

How could p_μ possibly be useful? We are hard-coding the coefficients of p_μ as the probabilities assigned by μ. So we have to use properties of p_μ other than their coefficients if we want to make use of our encoding.

What else might we use about a polynomial? The three ways of looking at a univariate polynomial $p \in \mathbb{R}[t]$ of degree d are:

(1) Its coefficients: $p(t) = a_0 + a_1 t + \cdots + a_d t^d$. (Note we write the lowest degree coefficients in this course first since then a_i corresponds to the probability of obtaining i.)

(2) Its zeros: $p(t) = a_d(t-\lambda_1)(t-\lambda_2)\ldots(t-\lambda_d)$. By the fundamental theorem of algebra, p has d (possibly complex) roots $\lambda_1, \ldots, \lambda_d \in \mathbb{C}$. Some may be equal.

(3) Its behavior as a function: $p(t) : \mathbb{C} \to \mathbb{C}$.

In this course we will learn **how these three views relate**.

Here, we're going to need to use (2) or (3) to prove unimodality (a property of the coefficients). Let's investigate its zeros. These are easy: the polynomial is already factored! The zeros are exactly $-\frac{1-p_i}{p_i}$ for $1 \le i \le d$. These are all, you might notice, real numbers.

This lets us apply Newton's inequalities, proved in 1707:

Lemma 1.1 (Newton's inequalities). *Let p be a real-rooted univariate polynomial of degree d with non-negative coefficients. Then, its coefficients are ultra-log concave, i.e., for any $1 \le i \le d-1$, we have:*

$$\left(\frac{a_i}{\binom{d}{i}}\right)^2 \ge \frac{a_{i-1}}{\binom{d}{i-1}} \cdot \frac{a_{i+1}}{\binom{d}{i+1}}$$

Furthermore, the set of positive indices form an interval of $\mathbb{Z}_{\ge 0}$.

Notice this is stronger than log concavity, which for a sequence of non-negative numbers $a_0, \ldots, a_d$ means that

$$a_i^2 \ge a_{i-1} \cdot a_{i+1}$$

for all $1 \le i \le d-1$.

Now let's see why this implies unimodality. By log concavity, $\frac{a_i}{a_{i-1}} \geq \frac{a_{i+1}}{a_i}$. In other words, the multiplicative rate of growth is decreasing. This is precisely why it's called log concavity.[1] To get unimodality, just notice that since the ratios are decreasing, while the ratio is above 1, the sequence increases. Once the ratio goes below 1, it must keep decreasing. We need that there are no "internal" zeros of this sequence, though. Otherwise, the log concavity condition becomes meaningless, as the ratios become undefined. For example, the sequence $1, 0, 0, 1$ is log concave but not unimodal: the sequence of ratios is 0, 0/0 (undefined) and then 1/0 (undefined).

In order to prove Newton's inequalities, we will need to understand real rooted polynomials better.

1.2 Closure Properties of Real Rooted Polynomials

Suppose $p \in \mathbb{R}[t]$ is a real rooted polynomial of degree d. Then:

(1) $p(ct)$ is real rooted for any $c \neq 0 \in \mathbb{R}$, as is $c \cdot p(t)$.

Proof. Let λ be a real root of p. Then λ/c is a real root of $p(ct)$. This implies $p(ct)$ has as many real roots as p, and as it has the same degree as $p(t)$, it must be real rooted as well. $c \cdot p(t)$ has the same roots as p. □

(2) $p'(t)$ is real rooted.

Proof. Let $\lambda_1 \leq \lambda_2 \leq \cdots \leq \lambda_{d'}$ be the distinct real roots of p. By Rolle's theorem,[2] there is a real root of p' between each pair of roots, giving us $d' - 1$ real roots. For each root with multiplicity $q > 1$, we get $q - 1$ additional roots, giving us the desired $d - 1$ real roots. □

(3) $\tilde{p}(t) = t^d p(1/t)$ is real rooted.

This reverses the coefficients. To give an example, $p(t) = 0 + \frac{1}{2}t + \frac{1}{2}t^2$ leads to $0t^2 + \frac{1}{2}t + \frac{1}{2}$.

Proof. We can write (where m is the number of zero roots):

$$p(t) = a_d t^m \prod_{j=m+1}^{d} (t - \lambda_j) = a_d t^d \prod_{j=m+1}^{d} (1 - \lambda_j/t)$$

but then

$$t^d p(1/t) = a_d \prod_{j=m+1}^{d} (1 - \lambda_j t),$$

which is a factorization into real roots, the ith equal to $1/\lambda_i$. □

[1] More formally, it's because the log of the sequence is a concave function. In particular, taking the log of both sides and dividing by two, we get:

$$\log(a_i) \geq \frac{\log(a_{i-1}) + \log(a_{i+1})}{2}$$

This is concavity: the midpoint is above the average of the endpoints.

[2] Recall this says that any real function f which is continuous and differentiable on $[a, c]$ and has $f(a) = f(c)$ has a point $a < b < c$ for which $f'(b) = 0$.

Let's now use these to prove Newton's inequalities.

Proof. Let $p = \sum_{j=0}^{d} a_j t^j$. We want to obtain, for each $1 \le i \le d-1$,

$$\left(\frac{a_i}{\binom{d}{i}}\right)^2 \ge \frac{a_{i-1}}{\binom{d}{i-1}} \cdot \frac{a_{i+1}}{\binom{d}{i+1}}$$

To do this, differentiate the polynomial $i-1$ times. This is real rooted by the closure properties, and removes the first $i-1$ coefficients $a_0, \dots, a_{i-2}$. Then, reverse the coefficients, and differentiate $d-i-1$ times. This removes the last $d-i-1$ coefficients of the original polynomial, $a_{i+2}, \dots, a_d$. What remains is a_{i-1}, a_i, a_{i+1}. What happened to each coefficient? First, each got differentiated $i-1$ times:

$$(i-1)!a_{i-1}, i!a_i t, \frac{(i+1)!}{2} a_{i+1} t^2.$$

The coefficients got reversed, so these terms were put at the end of the polynomial:

$$(i-1)!a_{i-1}t^{d+1-i}, i!a_i t^{d-i}, \frac{(i+1)!}{2} a_{i+1} t^{d-i-1}.$$

Finally, we differentiated $d-i-1$ times to get the final polynomial $p^*(t)$ of degree at most 2:

$$\begin{aligned}
p^*(t) &= (i-1)!\frac{(d-i+1)!}{2} a_{i-1}t^2 + i!(d-i)!a_i t + \frac{(i+1)!}{2}(d-i-1)!a_{i+1} \\
&= \frac{d!}{2}\left(\frac{(i-1)!(d-i+1)!}{d!} a_{i-1}t^2 + 2\frac{i!(d-i)!}{d!} a_i t + \frac{(i+1)!(d-i-1)!}{d!} a_{i+1}\right) \\
&= \frac{d!}{2}\left(\frac{a_{i-1}}{\binom{d}{i-1}} t^2 + 2\frac{a_i}{\binom{d}{i}} t + \frac{a_{i+1}}{\binom{d}{i+1}}\right)
\end{aligned}$$

Finally, let's use that this polynomial is real rooted. So we can apply the discriminant: remember, for $ax^2 + bx + c$, we must have $b^2 - 4ac \ge 0$ for the polynomial to be real rooted. Plugging this in after dividing by $d!/2$, which we can remove by (1):

$$\frac{4a_i^2}{\binom{d}{i}^2} \ge 4\frac{a_{i-1}}{\binom{d}{i-1}} \cdot \frac{a_{i+1}}{\binom{d}{i+1}}$$

Which is ultra-log concavity. We leave the fact that there are no internal zeros as an exercise. □

1.3 Poisson Binomial and Real Rootedness

We have seen that every Poisson Binomial distribution has a real rooted generating polynomial with non-negative coefficients. What about the other way around? Is it the case that every real rooted polynomial $p \in \mathbb{R}_{\ge 0}[t]$ with $p(1) = 1$ is the generating polynomial of a Poisson Binomial? It turns out the answer is yes!

Lemma 1.2. *Let $p \in \mathbb{R}_{\ge 0}[t]$ be a real rooted polynomial of degree d with $p(1) = 1$. Then, $p(t) = \prod_{i=1}^{d}(p_i t + (1-p_i))$ for some $p_1, \dots, p_d \in [0,1]$.*

Proof. Let $p(t) = \sum_{i=0}^{d} a_i t^i$ have roots $\lambda_1, \ldots, \lambda_d \leq 0$. Note that all roots must be non-positive, as the polynomial is non-zero and has no negative coefficients so $p(t) > 0$ for $t \geq 0$.

$$p(t) = a_d \prod_{i=1}^{d} (t - \lambda_i)$$

Now it's just a matter of rewriting this in the desired form. Note that since $p(1) = 1$, we have:

$$a_d = \frac{1}{\prod_{i=1}^{d} (1 - \lambda_i)}$$

Therefore,

$$p(t) = \frac{\prod_{i=1}^{d} (t - \lambda_i)}{\prod_{i=1}^{d} (1 - \lambda_i)} = \prod_{i=1}^{d} \left(\frac{1}{1 - \lambda_i} t - \frac{\lambda_i}{1 - \lambda_i}\right)$$

which is in the desired form if we set $p_i = \frac{1}{1-\lambda_i}$. □

So real rooted generating polynomials are quite simple. They are always Poisson Binomial distributions.

1.4 Beyond Univariate Polynomials

Things start to get more interesting when we encode distributions consisting of more than just one output, for example combinatorial distributions over a set of edges E, $\mu : 2^E \to [0, 1]$. Here, our generating polynomial has a variable for each element of E, so for example the distribution which is $\{e\}$ with probability $\frac{1}{2}$ and $\{e, f\}$ otherwise has the encoding $p_\mu(x_e, x_f) = \frac{1}{2} x_e + \frac{1}{2} x_e x_f$.

Of particular interest will be the uniform distribution over spanning trees, where $\mathcal{T}$ is the set of all spanning trees of a graph:

$$\frac{1}{|\mathcal{T}|} \sum_{T \in \mathcal{T}} \prod_{e \in T} x_e$$

And the matching polynomial, where M is the set of matched vertices:

$$\sum_{M \in \mathcal{M}} (-1)^{|M|} \prod_{v \in M} x_v$$

Notice that evaluated at i, this gives the number of matchings in the graph.

To recap: in this lecture, we studied univariate polynomials whose "zero-free region" was all of the complex plane excluding the real line. We were able to characterize precise properties of distributions whose generating polynomials had this zero free region. The above polynomials turn out to have a zero-free region as well, and we will be able to extract similarly useful characterizations of their underlying distributions.